

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26

**Understanding the Signal-to-noise Paradox in Decadal Climate Predictability from CMIP5  
and an Eddy-resolving GCM**

Wei Zhang<sup>1</sup>; Ben Kirtman<sup>1</sup>; Leo Siqueira<sup>1</sup>; Amy Clement<sup>1</sup>; Junfei Xia<sup>1</sup>

<sup>1</sup> Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, USA.

Corresponding author: Wei Zhang (wei.zhang@rsmas.miami.edu)

**Acknowledgments.** The research is supported by NOAA NA18OAR4310293, NA15OAR4320064, NSF OCE1419569, OCE1559151, and DOE DE-SC0019433. All the observations and CMIP5 model data used in this study are publicly available from the links and citations provided in the manuscript. CCSM4 model codes, experiments, and outputs were performed and archived at the University of Miami Center for Computational Science.

**Conflict of interest.** The authors declare that they have no conflict of interest.

27 **Abstract**

28 Recent research suggests the widespread existence of the signal-to-noise paradox in seasonal-to-decadal climate  
29 predictions. The essence of the paradox is that the signal-to-noise ratio in models can be unrealistically small and  
30 models may make better predictions of the observations than they predict themselves. The paradox highlights a  
31 potentially serious issue with model predictions as previous studies may underestimate the limit of predictability. The  
32 focus of this paper is two-fold: the first objective is to re-examine decadal predictability from the lens of the signal-  
33 to-noise paradox in the context of CMIP5 models. We demonstrate that decadal predictability is generally  
34 underestimated in CMIP5 models possibly due to the existence of the signal-to-noise paradox. Models underestimate  
35 decadal predictability in regions where it is likely for the paradox to exist. The second objective follows from the  
36 results of the first, attempting to determine if this underestimate of decadal predictability is, at least partially, due to  
37 missing ocean mesoscale processes and features in CMIP5 models. A suite of coupled model experiments is performed  
38 with eddy-resolving and eddy-parameterized ocean component. Compared with eddy-parameterized models, the  
39 paradox is less likely to exist in eddy-resolving models, particularly over eddy-rich regions. These also happen to be  
40 regions where increased decadal predictability is identified. We hypothesize that this enhanced predictability is due to  
41 the enhanced vertical connectivity in the ocean. The presence of mesoscale ocean features and associated vertical  
42 connectivity significantly influence decadal variability, predictability, and the signal-to-noise paradox.

43

44

45 Keywords: Signal-to-noise paradox; Decadal predictability; CMIP5; CCSM4; Eddy-resolving model; Vertical  
46 Connectivity

47

48

49

50

51

52

53

## 54 **1 Introduction**

55           There is a continuously growing demand for decadal climate predictions. Making skillful decadal predictions  
56 has potential benefits in terms of supporting decision-making processes in agriculture, energy and water management  
57 among other sectors (e.g., Kirtman et al. 2013; Kushnir et al. 2019; Merryfield et al. 2020). While seasonal climate  
58 prediction has matured into regular operational forecasts (e.g., Kirtman et al. 2014), forecasting the climate over  
59 decades has proven more challenging (Keenlyside et al. 2008; Meehl et al. 2014; Zhang and Kirtman 2019a).

60           One of the significant challenges in decadal prediction and often overlooked in previous studies is the so-  
61 called “signal-to-noise paradox” (e.g., Scaife et al. 2014; Siegert et al. 2016; Smith et al. 2019; Zhang and Kirtman  
62 2019b). The essence of the paradox is that the signal-to-noise ratio estimated in climate models can be too small.  
63 Specifically, models seem to be better at predicting observations than predicting themselves as the model ensemble  
64 mean forecasts are better correlated with observations than with individual ensemble members. Scaife et al. (2014)  
65 first discussed the signal-to-noise paradox in seasonal prediction of the winter North Atlantic Oscillation (NAO) index,  
66 and subsequently, a growing list of examples in different atmospheric and climate models has emerged (Scaife and  
67 Smith 2018). For example, Zhang and Kirtman (2019b, hereafter ZK19) developed a simple Markov model framework  
68 and provided a comprehensive assessment of the NAO index indicating the widespread existence of the signal-to-  
69 noise paradox in coupled models from the fifth Coupled Model Intercomparison Project (CMIP5). The Markov model  
70 framework can easily reproduce the signal-to-noise paradox, which is dependent on the magnitude of the persistence  
71 and noise variance. Smith et al. (2019) used multi-model decadal hindcasts from seven state-of-the-art coupled climate  
72 models with a total of 71 ensemble members suggesting the existence of the signal-to-noise paradox in decadal  
73 predictions. One of the key points highlighted in the Smith et al. (2019) paper was that model-based estimates of  
74 decadal predictability might actually be an underestimate, as previous studies could have misrepresented the noise, or  
75 underestimated the magnitude of the predictable signal due to limited ensemble size.

76           The specific examples of the paradox and the associated model errors noted above suggest that model based  
77 estimates of climate predictability may seriously underestimate the limit of predictability. As we begin to understand  
78 the mechanisms for the paradox, predictability estimates also need to be revisited. Where and to what extent is the  
79 paradox leading to substantial underestimates of the limit of predictability? The first goal of the paper is to re-examine  
80 decadal predictability from the lens of the signal-to-noise paradox in the context of the CMIP5 models.

81           The second goal of the paper follows from the results of the first. Essentially, the results from the first goal  
82 show that the CMIP5 models seriously underestimate the limit of decadal predictability, and we hypothesize in this  
83 second goal that this underestimate is, at least in part, due to missing ocean mesoscale processes and features in the  
84 CMIP5 models. Again, the results are presented in the context of the signal-to-noise paradox.

85           With the above in mind, there have been several studies examining the mechanisms for the paradox. For  
86 example, the signal-to-noise paradox has been attributed to a lack of persistence (Strommen and Palmer 2019; Zhang  
87 and Kirtman 2019b), weak extratropical air-sea coupling (Scaife and Smith 2018), stratospheric initialization (O'Reilly  
88 et al. 2019), and underestimated eddy feedbacks due to low atmospheric model resolution (Scaife et al. 2019). Little  
89 to no research, however, has asked how mesoscale ocean features affect the signal-to-noise paradox and associated  
90 estimate of decadal predictability. The role of ocean mesoscale processes is of particular interest since several previous  
91 studies have suggested that decadal SST variability in coupled models is improved when ocean mesoscale features  
92 and processes are resolved (e.g., He et al. 2018; Infanti and Kirtman 2019; Kim et al. 2018; Samanta et al. 2018;  
93 Kirtman et al. 2012, 2017; Siqueira and Kirtman 2016; Zhang and Kirtman 2019a; among others). As shown by  
94 Kirtman et al. (2017), for example, the resolved mesoscale ocean features can substantially influence large-scale  
95 climate variability, air-sea interactions, and predictability. Particularly in the North Atlantic region, a more realistic  
96 mean-state climate and improved representation of ocean-atmosphere coupling and decadal SST variability around  
97 the Gulf Stream region have been detected with eddy-resolving GCMs (Siqueira and Kirtman 2016). Given the  
98 importance of eddies on low-frequency variability and ocean-atmosphere coupling, the lack of ocean eddy resolution  
99 in current coupled models (e.g., eddy-parameterized models used in CMIP5) can potentially affect the estimates of  
100 decadal climate predictability.

101           In addition to the overall representation of decadal variability, the second goal of this study is motivated by  
102 the hypothesis that low-resolution eddy-parameterized GCMs may misrepresent or even lack the vertical  
103 communication in the subsurface to the deeper ocean, contributing to a lack of persistence in models and thus the  
104 signal-to-noise paradox. The underestimated vertical communication between the deep ocean and surface processes  
105 in CMIP5 models compared to observations has been recently explored by Kravtsov (2020). Kravtsov (2020)  
106 introduced an updated linear energy-balance model considering the heat exchange between ocean mixed layer and  
107 thermocline in the Atlantic and Pacific oceans. By fitting the observed and CMIP5 model-simulated SST with the  
108 energy-balance model, Kravtsov (2020) identified stronger vertical communication between the deep ocean and

109 surface processes in observations than CMIP5 models, contributing to a larger fraction of predictable variability at  
110 decadal timescales. This significant difference in decadal potential predictability between observations and CMIP5  
111 models, as suggested by Kravtsov (2020), may lead to the signal-to-noise paradox. We note that the time-scales  
112 associated with the vertical connectivity arguments presented here are considerably shorter than the time-scales  
113 addressed in Clement et al. (2015) who question the role of sub-surface ocean dynamics in the Atlantic multi-decadal  
114 variability based on coarse resolution CMIP3 and CMIP5 models.

115 In this study, we first examine the decadal potential predictability in observations and CMIP5 models from  
116 a diagnostic perspective, i.e., the first goal. Again, through the lens of the signal-to-noise paradox we use the Markov  
117 model framework developed in ZK19 to diagnose predictability. In terms of the second goal, distinct from Kravtsov  
118 (2020) who estimated the coupling parameters between thermocline and mixed layer in the energy-balance model, we  
119 perform a suite of model experiments with and without resolved ocean mesoscale features, again through the lens of  
120 the paradox. We argue that high-resolution models with resolved ocean mesoscale features have stronger vertical  
121 connectivity in the subsurface to the deeper ocean than low-resolution models, which may potentially, or at least  
122 partially eliminate the signal-to-noise issue and thus improve decadal predictability over decadal timescales.

123

## 124 **2 Data and Method**

### 125 **2.1 Observations and CMIP5 models**

126 Three observational monthly sea surface temperature (SST) datasets are used in this study; namely, the  
127 National Oceanic and Atmospheric Administration (NOAA) Extended Reconstructed SST version-5 (ERSST; Huang  
128 et al. 2017) on  $2^\circ \times 2^\circ$  grids for 1854-present, the Hadley Center Global Sea Ice and SST data set (HadISST; Rayner  
129 et al. 2003) from 1870 to 2017 with a spatial resolution of  $1^\circ \times 1^\circ$ , and the Centennial Observation-Based Estimates of  
130 SST version-2 (COBE; Hirahara et al. 2014) from 1850 to 2017 on the same grid as HadISST data. Monthly mean sea  
131 level pressure (SLP) data is obtained from three resources, including two 20th century reanalysis datasets from the  
132 NOAA (20CR;  $2^\circ \times 2^\circ$ ; 1871-2012; Compo et al. 2011) and the European Centre for Medium-Range Weather Forecasts  
133 (ERA20C;  $1^\circ \times 1^\circ$ ; 1900-2010; Poli et al. 2016), as well as the Hadley Centre's Mean SLP data (HadSLP;  $5^\circ \times 5^\circ$ ; 1850-  
134 2004; Allan and Ansell 2006).

135 Both the historical (HIST, first realization) and the preindustrial control (PI) simulations of thirty CMIP5  
136 models are used in this study to compare with observations (Table 1). We only use the first realization (r1i1p1) of

137 each CMIP5 model to equally weight each model in the multi-model mean estimates. The HIST simulations are  
 138 simulations of recent past climate (1850-2005) forced by changing conditions, while the PI simulations are  
 139 preindustrial coupled ocean-atmosphere control simulations with non-evolving preindustrial conditions (Taylor et al.  
 140 2012). Variability in the PI simulations is generated only through interactions internal to the coupled system, while  
 141 variability in the HIST simulations is also due to natural and anthropogenic forcing (Murphy et al. 2017).

	<b>Models</b>	<b>Atmospheric Resolution</b>	<b>PI Length (years)</b>	<b>Data/Modeling Institute</b>
1	ACCESS1-0		500	Commonwealth Scientific and Industrial Research Organization and Bureau of Meteorology (Australia)
2	ACCESS1-3	145 × 192	500	
3	BCC-CSM1-1	64 × 128	500	China Meteorological Administration
4	BCC-CSM1-1-m	160 × 320	400	
5	BNU-ESM	64 × 128	559	Beijing Normal University
6	CanESM2	64 × 128	996	Canadian Centre for Climate Modeling and Analysis National Center for Atmospheric Research
7	CCSM4	192 × 288	1051	
8	CESM1-BGC		500	National Science Foundation and Department of Energy
9	CESM1-CAM5	192 × 288	319	
10	CESM1-FASTCHEM		222	
11	CMCC-CM	240 × 480	330	Centro Euro-Mediterraneo per I Cambiamenti Climatici
12	CMCC-CMS	96 × 192	500	
13	CNRM-CM5	128 × 256	850	Centre National de Recherches Meteorologiques and Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique
14	CSIRO-Mk3-6-0	96 × 192	500	Australian Commonwealth Scientific and Industrial Research Organization and Queensland Climate Change Centre of Excellence
15	GFDL-CM3		500	Geophysical Fluid Dynamics Laboratory
16	GFDL-ESM2G	90 × 144	500	
17	GFDL-ESM2M		500	
18	GISS-E2-H-CC		251	Goddard Institute for Space Studies
19	GISS-E2-R-CC	90 × 144	251	
20	HadGEM2-CC		240	Met Office Hadley Centre
21	HadGEM2-ES	145 × 192	576	
22	inmcm4	120 × 180	500	Institute for Numerical Mathematics
23	IPSL-CM5A-LR	96 × 96	1000	Institut Pierre-Simon Laplace
24	IPSL-CM5A-MR	143 × 144	300	
25	IPSL-CM5B-LR	96 × 96	300	
26	MPI-ESM-LR		1000	Max Planck Institute for Meteorology
27	MPI-ESM-MR	96 × 192	1000	
28	MPI-ESM-P		1156	
29	NorESM1-ME		252	Norwegian Climate Centre
30	NorESM1-M	96 × 144	501	

142  
 143 Table 1. CMIP5 models used in this study. PI accounts for the preindustrial control simulations. We also use all the  
 144 historical simulations of CMIP5 models from 1870 to 2005. The CMIP5 model outputs and associated descriptions  
 145 can be found in the CMIP5 archive (<http://cmip-pcmdi.llnl.gov/cmip5>).

146

## 147 **2.2 CCSM4 model experiments**

148 A suite of model experiments is performed with the National Center for Atmospheric Research Community  
149 Climate System Model Version 4 (CCSM4; see overview in Gent et al. 2011). In the low-resolution eddy-  
150 parameterized experiment (hereafter referred to as LR), we use 1° atmosphere (CAM4) and land (CLM3.5) models  
151 coupled to the ocean (POP2) and sea-ice (CICE4) models with 1° resolution; all the component models are linked and  
152 coordinated through the CCSM flux coupler. We conduct the LR CCSM4 experiment with present-day forcing (e.g.,  
153 greenhouse gas concentrations from 1990) for 500 years, and the first 200 years are viewed as a spin-up period.  
154 Different from the LR experiment, we employ a 0.5° atmospheric model coupled to 0.1° ocean and sea ice component  
155 models (Kirtman et al. 2012) in the high-resolution eddy-resolving experiment (hereafter referred as HR). We first  
156 perform a 155-year standard control simulation with the same greenhouse gas concentrations as in the LR experiment,  
157 and the first 100 years of the simulation are taken as spin-up and are discarded in the analysis. Restarting from the  
158 first experiment with small perturbations, we run two other experiments for 70 years, each with the first 20 years taken  
159 as spin-up periods. In total, we analyze here 155 years of HR simulations and 300 years of LR simulations.

160

## 161 **2.3 Markov model framework**

162 The Markov model framework is extensively described in ZK19 and has proven useful to determine the  
163 existence of the signal-to-noise paradox. The design of the Markov model framework starts from a linear signal-plus-  
164 noise model assuming that the future state forecasts depend linearly on the current state predictor and a stochastic  
165 noise term. The observations and models can be simulated with a statistical Markov chain model:

$$O^{n+1} = \alpha O^n + N \quad (1)$$

$$M_i^{n+1} = \beta O^n + P_i \quad (2)$$

$$\overline{M}^{n+1} = \beta O^n + \overline{P} \quad (3)$$

166 where  $\{O\}$  is the observation, and  $\{M\}$  is the model forecasts initialized with observations and perturbed with different  
167 noise realizations  $\{P_i\}$  ( $i = 1, 2, 3, \dots$ ).  $\alpha$  and  $\beta$  are estimated as the lag-1 autocorrelation coefficients implying the  
168 persistence of the system.  $N$  and  $P$  are simply modeled as white-noise processes (see also Kirtman et al. 2005).  $\{\overline{M}\}$   
169 and  $\{\overline{P}\}$  are model ensemble mean forecasts and ensemble mean noise. Following the procedure in ZK19, we can  
170 analytically derive the correlation between the model ensemble mean forecasts and observations ( $\text{corr}(\overline{M}, O)$ ) and the

171 correlation between the model ensemble mean forecasts and individual ensemble members ( $corr(\bar{M}, M_i)$ ), and thus  
 172 the ratio of squared correlation ( $RSC$ ):

$$RSC = \frac{corr^2(\bar{M}, O)}{corr^2(\bar{M}, M_i)} = \frac{\alpha^2 \beta^2 \sigma_N^2 + \alpha^2 (1 - \alpha^2) \sigma_P^2}{\beta^2 \sigma_N^2} \quad (4)$$

173 where  $\sigma_N^2$  and  $\sigma_P^2$  are noise variance for the observation and model ensemble members, which can be estimated as the  
 174 total variance in observations and model forecasts multiplied by a factor of  $1 - \alpha^2$  and  $1 - \beta^2$ , respectively. Based  
 175 on Scaife and Smith (2018), there can be a signal-to-noise paradox when the  $RSC$  is greater than 1.0. So (4) becomes,

$$\frac{\alpha^2}{\sigma_N^2} > \frac{\beta^2}{\sigma_P^2} \quad (5)$$

176 Distinct from ZK19, who applied the Markov model framework to the monthly NAO index, this study further  
 177 examines the signal-to-noise paradox, especially in SST fields based on CMIP5 HIST and PI simulations with a focus  
 178 on decadal timescales.

179

## 180 **3 Results and Discussion**

### 181 **3.1 Underestimated decadal climate predictability in CMIP5 models**

182 Decadal climate predictability is quantified using the potential predictability variance ratio (Boer 2004;  
 183 Zhang et al. 2017), which indicates the relative intensity of decadal variability and provides an efficient approach to  
 184 estimate decadal predictability from a diagnostic perspective. As suggested by Boer (2004), the total climate variability  
 185 ( $\sigma_{Tot}^2$ ) can be decomposed into a low-frequency component ( $\sigma_{Lf}^2$ ) that may be potentially predictable and a high-  
 186 frequency unpredictable noise component ( $\sigma_{Err}^2$ ). This approach assumes that the slower potentially predictable  
 187 component is independent from the noise term ( $\sigma_{Lf}^2 = \sigma_{Tot}^2 - \sigma_{Err}^2$ ). Therefore, decadal potential predictability can  
 188 be defined as the ratio of decadal-scale variability with respect to the total variability ( $\sigma_{Lf}^2 / \sigma_{Tot}^2 \times 100\%$ ). In this  
 189 study, we extract decadal-scale climate variability by applying a 5-year low-pass Butterworth filter, after subtracting  
 190 the annual cycle and linearly detrending the monthly time series from each observation and model simulation. The  
 191 observational mean (based on ERSSTv5, COBE-SST2, and HadISST) and multi-model mean estimates (based on  
 192 thirty CMIP5 model HIST simulations) of decadal SST predictability are shown in Figs. 1a and 1b. Note that SST in  
 193 model simulations is obtained by masking all the non-ocean regions of surface temperature, and for each observation  
 194 and model simulation, SST is bilinearly interpolated to  $2^\circ \times 2^\circ$  grids before analysis. In observations, considerable

195 ocean regions display relatively higher values of decadal SST predictability such as the North Atlantic, Western Pacific,  
196 Tropical Indian Ocean, and Southern Ocean, which is generally consistent with several earlier studies (e.g., Ding et  
197 al. 2016; Zhang and Kirtman 2019a). Meanwhile, relatively lower decadal SST predictability is detected over the  
198 Eastern Tropical Pacific, with values ranging from 25% to 60% gradually increasing westward, implying that decadal  
199 climate in this region may be still potentially predictable to some extent, though the Eastern Tropical Pacific is  
200 dominated by the El Niño–Southern Oscillation (ENSO) at interannual timescales. Uncertainty remains in the long-  
201 term predictability over the Eastern Tropical Pacific region (e.g., Gonzalez and Goddard 2016. Kirtman and Schopf  
202 1998; Kravtsov 2012; Newman 2007), and as suggested by Wittenberg et al. (2014), for example, the ENSO potential  
203 predictability is lost after the 3-4 year range in the absence of external forcing.

204           Compared with observational estimates, we find much lower values of decadal SST predictability for the  
205 CMIP5 multi-model mean estimates, except some regions such as the Northeastern Pacific and subpolar North Atlantic  
206 (Fig. 1b). High decadal SST predictability in the subpolar North Atlantic is identified, and consistent with the results  
207 based on observations, where decadal SST variability is significantly impacted by internal atmospheric noise and  
208 subsurface ocean dynamics (Zhang and Kirtman 2019a). Fig. 1c shows the difference between the observational and  
209 CMIP5 multi-model estimates in six different ocean regions, namely, the North Atlantic (NA; 20-65°N, 80-0°W),  
210 North Pacific (NP; 20-60°N, 120°E-120°W), Eastern Tropical Pacific (ETP; 10°S-10°N, 180-85°W), Tropical  
211 Atlantic (TA; 15°S-15°N, 80-0°W), Tropical Indian (TI; 15°S-25°N, 40-100°E), and Southern Ocean (SO; 65-40°S,  
212 0-360°E). The observational estimates show higher decadal SST predictability than most of the CMIP5 models except  
213 in the North Pacific; that is, decadal SST predictability is generally underestimated in CMIP5 models. The spatial  
214 distribution of decadal SST predictability based on CMIP5 models in the North Pacific is substantially different from  
215 those based on observational estimates, with larger values concentrated in the subpolar gyre, despite that the mean  
216 estimates in models and observations are somewhat comparable. Notably, there is significant inconsistency among  
217 different observational estimates, especially in the Tropical Atlantic and the Southern Ocean; also, thirty CMIP5  
218 models used here show better agreement in the North Atlantic and North Pacific than other regions.

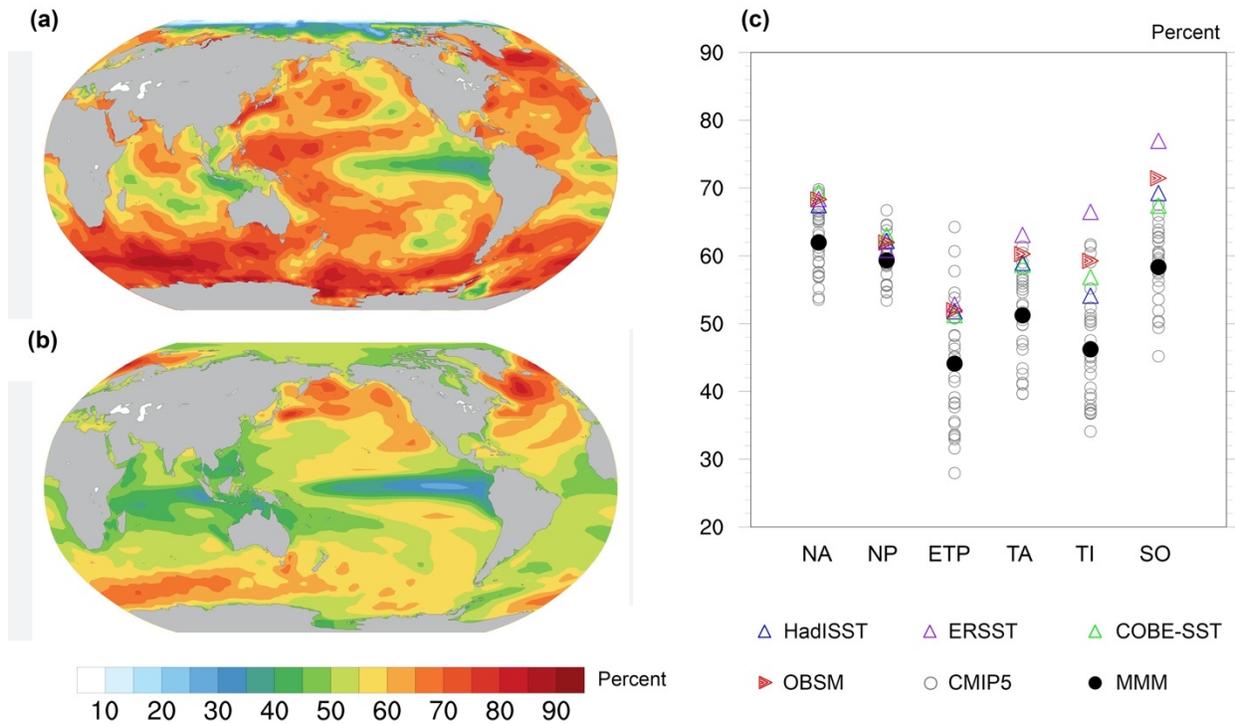
219           The impact of linear detrending is assessed in Fig. 2, which shows decadal SST predictability without  
220 detrending and with the regression-detrending method following Ting et al. (2009) by regressing out the global mean  
221 SST warming trend. Similar spatial patterns of decadal SST predictability is found with the different detrending  
222 methods and even without detrending, supporting our argument of the underestimated decadal SST predictability in

223 CMIP5 models.

224 Fig. 3 is the same as Fig. 1 but for decadal SLP predictability. The observations (ERA20C, 20CR, and  
225 HadSLP) present higher values of decadal SLP predictability than almost all the CMIP5 models, indicating that CMIP5  
226 models also underestimate the observed decadal SLP predictability. The underestimated decadal predictability in  
227 models is not merely due to the underestimate of decadal-scale SST and SLP variance in CMIP5 models. For example,  
228 the variance can be comparable or even higher than the observed variance estimate in substantial regions (Fig. 4). The  
229 results shown here suggest that the CMIP5 models largely underestimate the decadal predictability, but not necessarily  
230 the decadal variance.

231 Wang et al. (2015) diagnose the leading EOF modes of SST on monthly and decadal timescales and argue  
232 that compared with observational estimates, both CMIP3 and CMIP5 models fail to capture the accurate spatial  
233 structure of SST variability. The striking disagreement in observational estimates and CMIP5 models may stem from  
234 the ocean-atmosphere coupling (Li et al. 2013; Sun et al. 2015), ocean dynamics (Kirtman et al. 2012), and intrinsic  
235 model errors (Gupta et al. 2013; Richter 2015), which requires further investigation.

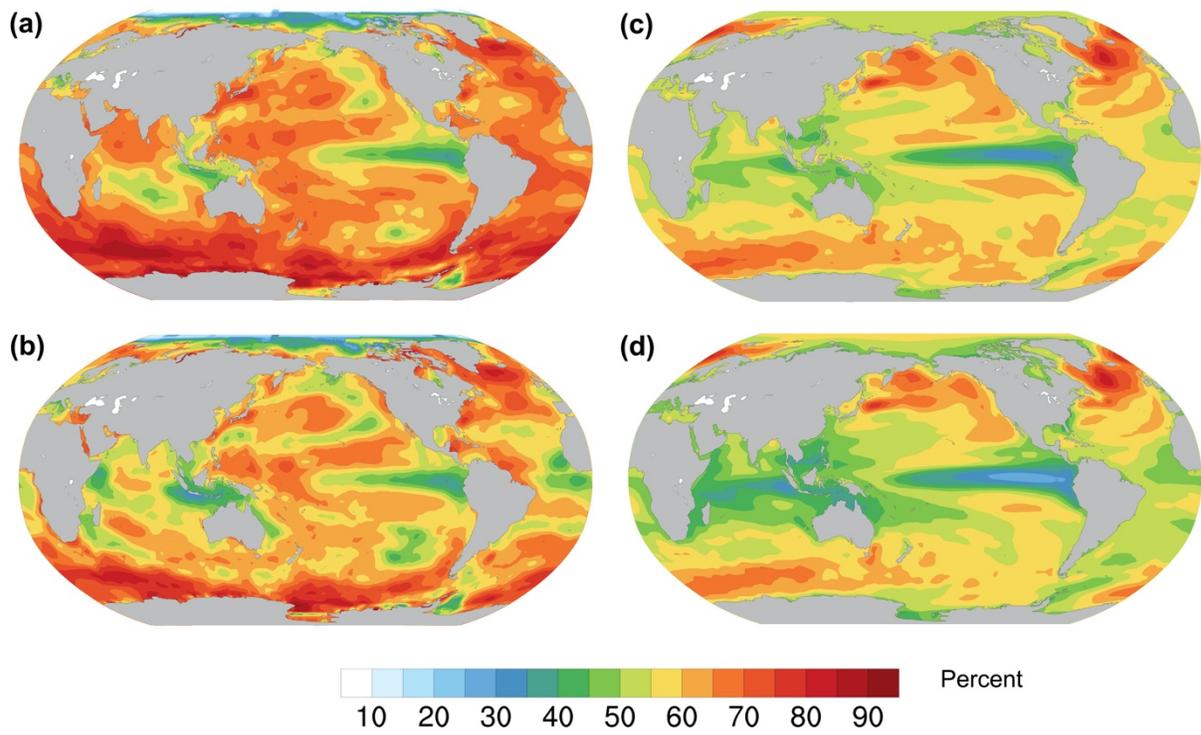
236



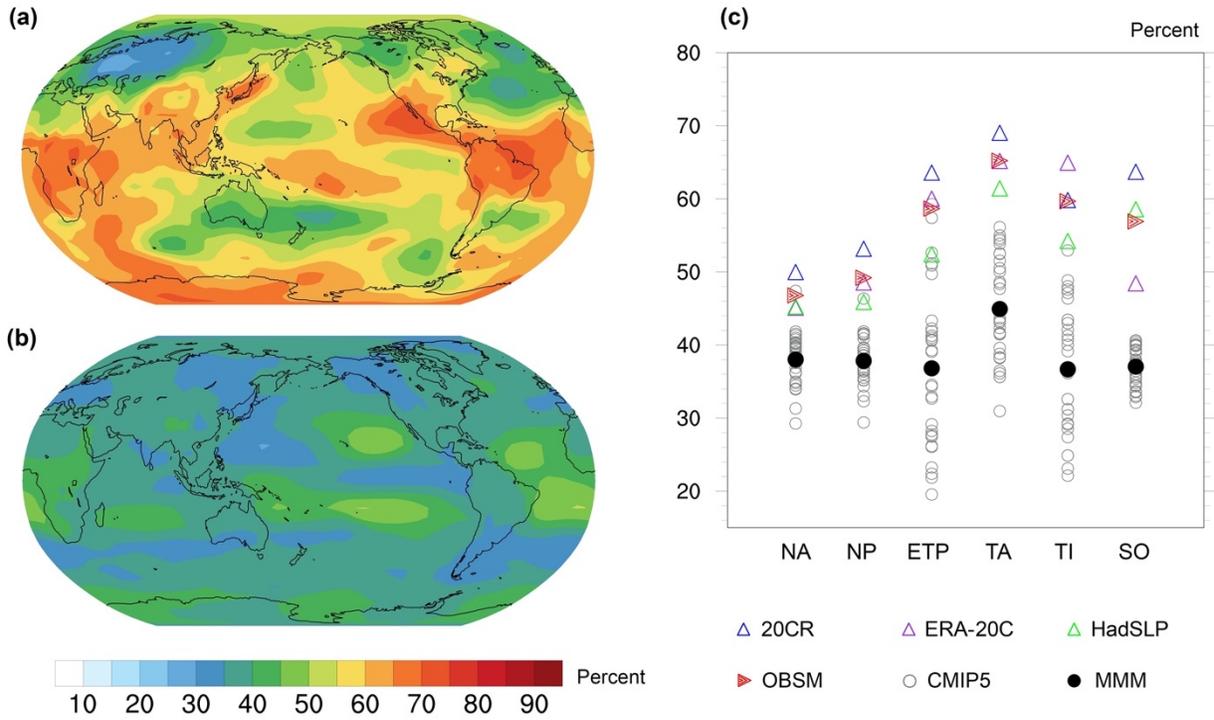
237

238 Fig. 1. Decadal SST predictability based on observations and CMIP5 models. (a) Observational mean estimates based  
239 on three observational SST datasets. We calculate decadal SST predictability for each observational dataset and then

240 take the average as the observational mean estimates. (b) Multi-model mean estimates based on 30 CMIP5 historical  
 241 simulations. (c) Difference of observational and multi-model mean estimates in six different ocean regions, namely,  
 242 the North Atlantic (NA; 20-65°N, 80-0°W), North Pacific (NP; 20-60°N, 120°E-120°W), Eastern Tropical Pacific  
 243 (ETP; 10°S-10°N, 180-85°W), Tropical Atlantic (TA; 15°S-15°N, 80-0°W), Tropical Indian (TI; 15°S-25°N, 40-  
 244 100°E), and Southern Ocean (SO; 65-40°S, 0-360°E).  
 245 HadISST = Hadley Centre Sea Ice and Sea Surface Temperature. ERSST = Extended Reconstructed Sea Surface  
 246 Temperature. COBE-SST = Centennial in situ Observation-Based Estimates. OBSM = observational mean estimates.  
 247 CMIP5 = Coupled Model Intercomparison Project 5. MMM = Multi-model Mean estimates.  
 248



249  
 250 Fig. 2. Impact of linear detrending on decadal SST predictability. (a) Observational mean estimates of decadal SST  
 251 predictability without any detrending. (b) Observational mean estimates of decadal SST predictability using the  
 252 regression-detrending method (regressing out the global mean warming trend). (c) Multi-model mean estimates of  
 253 decadal SST predictability without any detrending. (d) Multi-model mean estimates of decadal SST predictability  
 254 using the regression-detrending method.  
 255

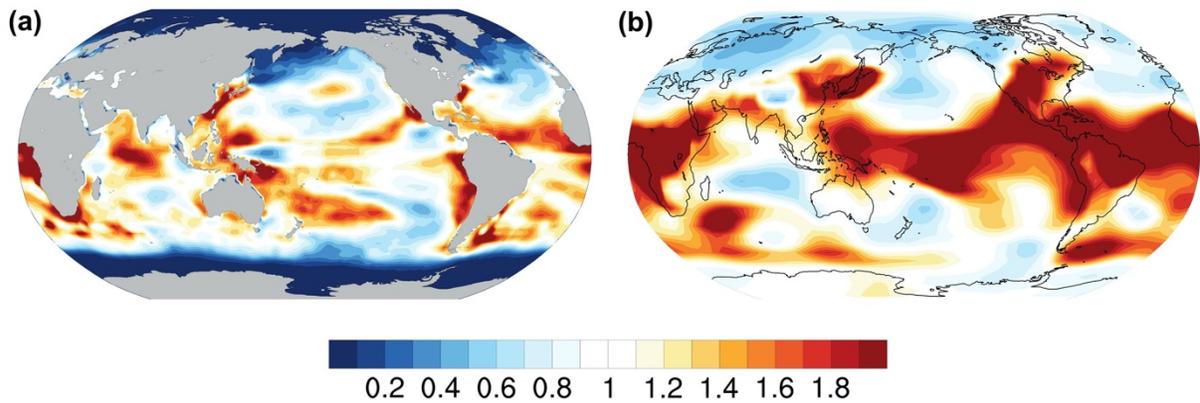


256

257 Fig. 3. Same as Fig. 1 but for decadal SLP predictability. (a) Observational mean estimates based on three  
 258 observational SLP datasets. (b) Multi-model mean estimates based on 30 CMIP5 historical simulations. (c) Difference  
 259 of observational and multi-model mean estimates in six different ocean regions.

260 20CR = NOAA-CIRES-DOE Twentieth Century Reanalysis. ERA-20C = ECMWF's Atmospheric Reanalysis of the  
 261 Twentieth Century. HadSLP = Hadley Centre's Mean Sea Level Pressure.

262



263

264 Figure 4. Variance ratio between observations and CMIP5 models for (a) decadal SST variability and (b) decadal SLP  
 265 variability. SST in models is obtained by masking all the non-ocean regions of surface temperature. Low ratio of  
 266 decadal SST variance between models and observations in polar regions is significantly affected by sea ice temperature.

### 267 **3.2 Signal-to-noise paradox in decadal climate predictability**

268           Recent research reveals the widespread existence of the signal-to-noise paradox in climate models. Here we  
269 ask is the underestimated decadal predictability in CMIP5 models related to the signal-to-noise paradox? Fig. 5  
270 encapsulates the relationship between the paradox and predictability. First, Fig. 5a shows the difference in decadal  
271 SST predictability between observations and CMIP5 HIST simulations, which is indicative of the underestimation in  
272 CMIP5 models. Based on the Markov model framework (see methods in Section 2c), the chance of existence for the  
273 signal-to-noise paradox based on thirty CMIP5 HIST simulations is examined in the low-pass filtered SST field (Fig.  
274 5b). Each SST observation and model simulation is detrended and normalized before analysis. The assessed likelihood  
275 of the existence of the paradox is designed following the Fifth Assessment Report of the United Nations  
276 Intergovernmental Panel on Climate Change (IPCC 2014): very likely 90-100%, likely 66-90%, about as likely as not  
277 33-66%, and very unlikely 0-33%. The patterns in the possibility of the existence of the signal-to-noise paradox (Fig.  
278 5b) is somewhat consistent with those in Fig. 5a, implying a possible relationship between the underestimated decadal  
279 SST predictability and the signal-to-noise paradox. Models are likely to underestimate decadal predictability in regions  
280 where it is likely to have a signal-to-noise paradox, especially around the Tropical Atlantic and the Tropical Indian  
281 Ocean and eddy-rich regions, including the Gulf Stream, the Kuroshio Current, and the Southern Ocean.

282           This relationship between the paradox and predictability is further supported by the ratio of squared  
283 correlation over six ocean areas (same as Fig. 1c) estimated for each model in CMIP5 (Fig. 6). The regional mean  
284 SST index for each ocean region is created and then assessed with the ratio of squared correlation for each model.  
285 There would be a signal-to-noise paradox when the ratio of squared correlation is higher than 1.0, as suggested by  
286 Zhang and Kirtman (2019b). Here we show that the signal-to-noise paradox is very likely to occur in extratropical  
287 regions (e.g., the North Atlantic), the Tropical Atlantic and the Southern Ocean; meanwhile, only about half of the  
288 CMIP5 models used in this study indicate a paradox in the North and Eastern Tropical Pacific.

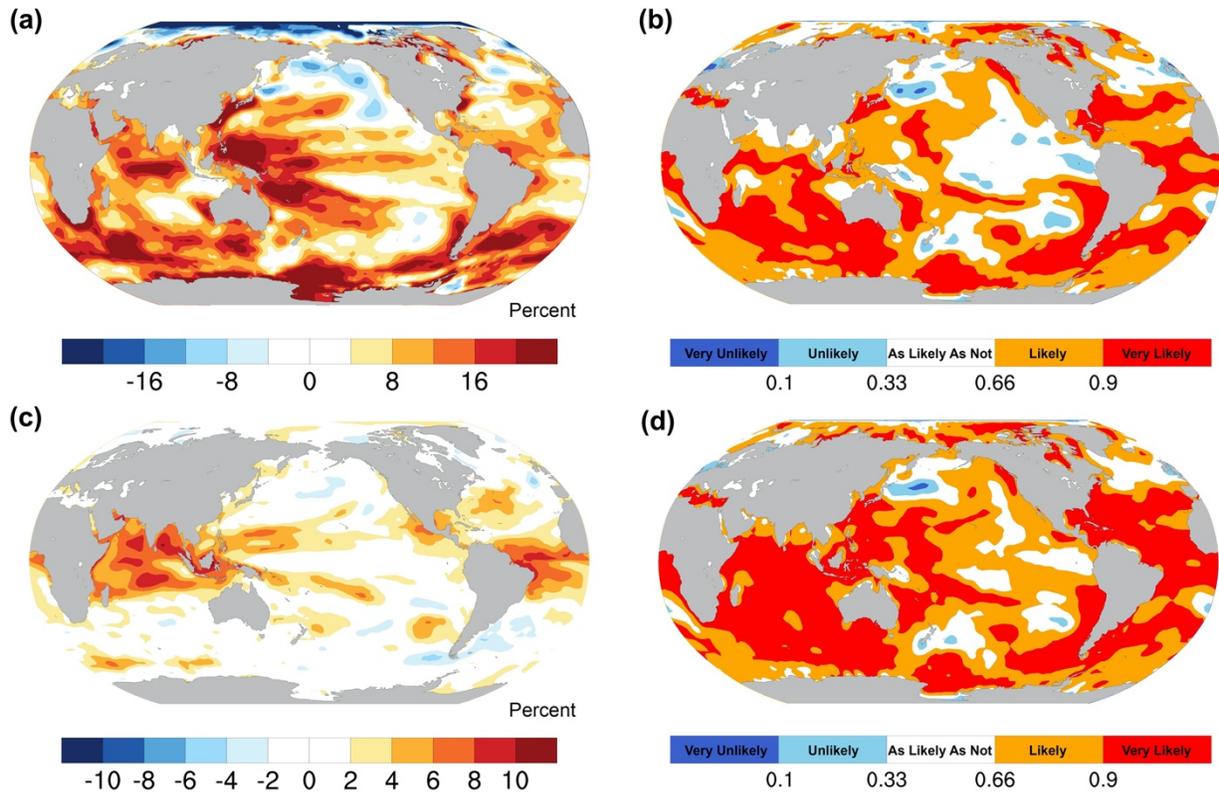
289           The distributions for the chance of occurrence of the signal-to-noise paradox in this study show some  
290 consistency with several previous studies (Eade et al. 2014; Smith et al. 2019); for instance, Eade et al. (2014) show  
291 the distribution of the signal-to-noise paradox in surface air temperature by the ratio of predictable component based  
292 on a multi-model ensemble of decadal hindcasts from the Met Office decadal prediction system (DePreSys; Knight et  
293 al. 2014) and four CMIP5 models.

294           The impact of external forcing on decadal SST predictability and the signal-to-noise paradox is also examined

295 here (Figs. 5c and 5d). Fig. 5c displays the difference of decadal SST variability between thirty CMIP5 HIST and PI  
296 simulations, with the latter having constant external forcing. The most noticeable difference between CMIP5 HIST  
297 and PI simulations appears in the tropics such as the Tropical Atlantic and the Tropical Indian ocean as longer decadal  
298 SST predictability is found in these regions, accompanied by a higher chance of the signal-to-noise paradox.  
299 Consistent with previous studies (e.g., Goddard et al. 2013; Guemas et al. 2013; Meehl et al. 2014), the Tropical Indian  
300 ocean stands out as the area significantly affected by the externally forced variability, which is shown to be much  
301 larger than the internally generated variability in both uninitialized simulations and initialized decadal hindcasts. It is  
302 worth mentioning that decadal SST predictability in CMIP5 models is smaller than the observational estimates, even  
303 in HIST simulations with evolving external forcing. We argue that CMIP5 models may underestimate the externally  
304 forced trend in the Tropical Indian ocean, though we cannot exclude the role of internal dynamics and any other  
305 associated factors in decadal predictability.

306 The Tropical Atlantic is another region of emerging interest in near-term climate predictability, where  
307 external forcing acts as an important factor driving decadal variability (Yeager and Robson 2017). The Tropical  
308 Atlantic has long been considered as a region with significant SST bias and poor upper ocean thermal structure and  
309 limited decadal predictability (Harlaß et al. 2018; Patricola et al. 2012; Xu et al. 2014). Shaffry et al. (2017) utilized  
310 a high-resolution eddy-permitting coupled GCM (ocean model resolution  $1/3^\circ \times 1/3^\circ$ ) and showed improved decadal  
311 prediction skills compared with low-resolution models, especially over the Tropical Atlantic region, pointing toward  
312 the importance of model resolution.

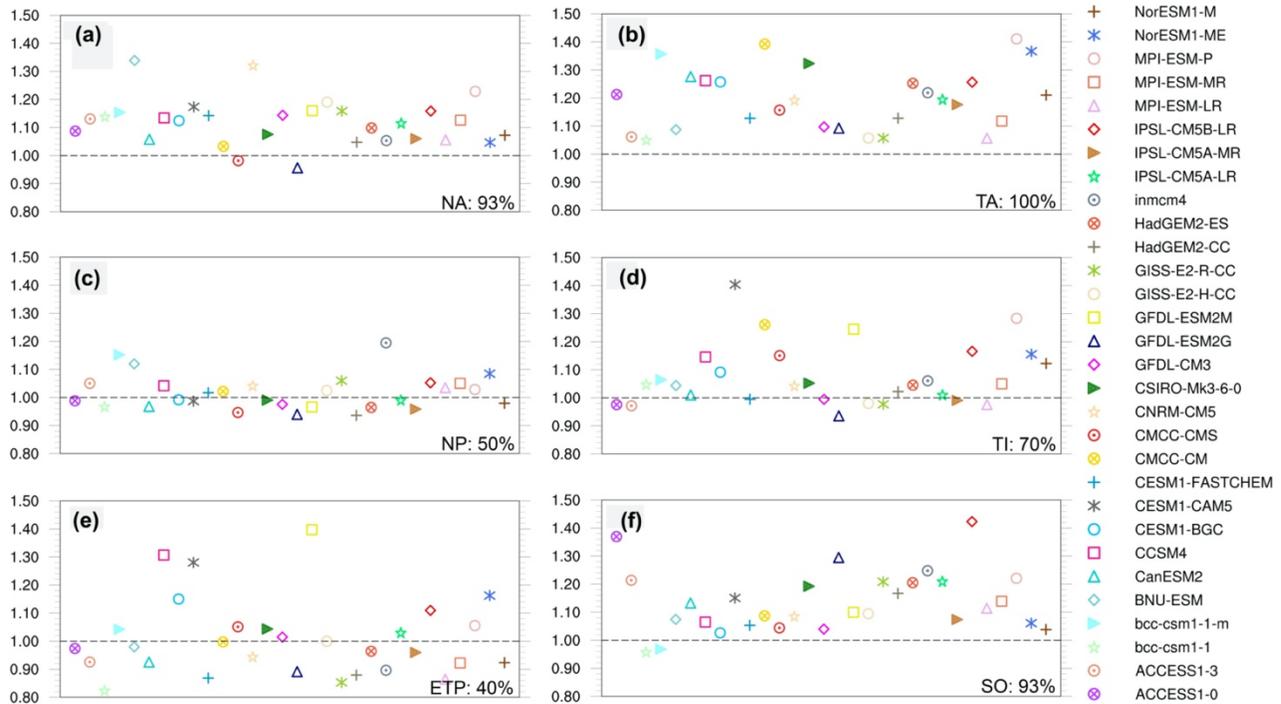
313 Furthermore, the coexistence of the underestimated decadal SST predictability and the high chance of  
314 occurrence for the signal-to-noise paradox in eddy-rich regions is suggestive of the lack of ocean model resolution in  
315 CMIP5 models. This is possible because all the coupled models in CMIP5 use eddy parameterized ocean models that  
316 may have weak vertical connectivity between ocean mixed layer and thermocline (Kravtsov 2020). The role of  
317 mesoscale ocean eddies and fronts, particularly in the western boundary regions has been highlighted in previous work  
318 (Bryan et al. 2010; Kirtman et al. 2012; Minobe et al. 2008; Siqueira and Kirtman 2016), and is a potential source of  
319 decadal predictability that has not been fully accounted for or leveraged.



320

321 Fig. 5. Existence of the signal-to-noise paradox in CMIP5 models and the impact of external forcing. (a) Difference  
 322 of decadal SST predictability between observations and CMIP5 historical simulations. (b) Chance of existence for the  
 323 signal-to-noise paradox based on 30 CMIP5 historical simulations. Each SST simulation is detrended and normalized  
 324 before analysis. The existence of the signal-to-noise paradox is estimated based on the Markov model framework. (c)  
 325 Difference of decadal SST predictability between CMIP5 historical and pre-industrial control simulations, suggesting  
 326 the impact of external forcing. (d) Chance of existence for the signal-to-noise paradox based on 30 CMIP5 pre-  
 327 industrial control simulations.

328



329

330 Fig. 6. Ratio of squared correlation estimated based on the Markov model framework for 30 CMIP5 models (historical  
 331 simulations) in six different ocean regions, including (a) NA: North Atlantic, (b) NP: North Pacific, (c) ETP: Eastern  
 332 Tropical Pacific, (d) TA: Tropical Atlantic, (e) TI: Tropical Indian, and (f) SO: Southern Ocean.

333

### 334 3.3 Advancing decadal predictability from an eddy-resolving GCM

335

336

337

338

339

340

341

342

343

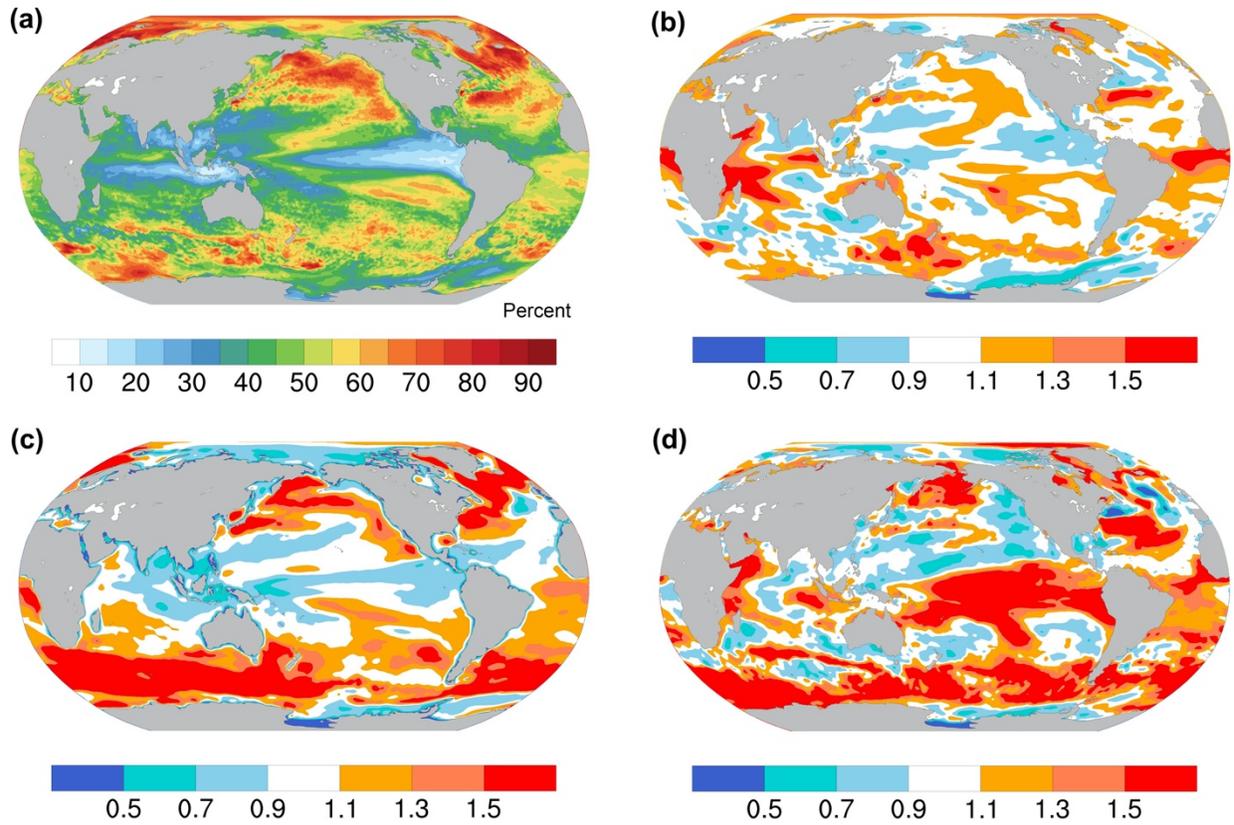
344

The following is based on the hypothesis that the presence of ocean mesoscale processes and features and the associated vertical connectivity affects decadal variability, predictability, and the signal-to-noise paradox. Specifically, coupled models with eddy-resolving component may, at least partially, reduce the signal-to-noise issue and thus improve decadal-scale climate predictability. The enhanced predictability we argue, is in part, due to the enhanced vertical connectivity in the ocean. This enhanced vertical connectivity allows the deeper ocean to more efficiently communicate with the surface, which, given the slower sub-surface time-scales, leads to longer surface predictability. To test this argument, we perform a suite of model experiments using CCSM4 with HR (eddy-resolved;  $0.1^\circ \times 0.1^\circ$ ) and LR (eddy-parameterized;  $1^\circ \times 1^\circ$ ) ocean component models. The details of the CCSM4 model setup and experiment design have been provided in Section 2b, which are generally consistent with Kirtman et al. (2017), but here we employ a much longer LR simulations.

345 Fig. 7 encapsulates how resolved ocean eddies affect decadal predictability estimates. In particular, Figs. 7a  
346 and 7b show the global distributions of decadal SST potential predictability based on CCSM4 HR and the difference  
347 of decadal SST predictability between HR and LR simulations, respectively. In the HR simulations, we find relatively  
348 higher decadal SST predictability in eddy-rich regions such as the Gulf Stream and Kuroshio Current systems, Tropical  
349 Atlantic, and Southern Ocean, where decadal SST variability is also increased (Fig. 7c). Perhaps surprising is that we  
350 also detect higher decadal SST predictability in HR over the Tropical Indian and the Tropical Atlantic regions,  
351 suggesting the strong influence of mesoscale ocean features on decadal SST predictability in addition to the impact of  
352 external forcing (e.g., Guemas et al. 2013; Meehl et al. 2014). We also note that decadal SST variability is remarkably  
353 elevated with HR simulations over western and eastern boundary current regions in the extratropics (Fig. 7c).

354 In the subpolar North Atlantic, both the LR and HR simulations show relatively longer decadal SST  
355 predictability (Fig. 7a), which is closely related with the persistence of low-frequency ocean heat content (OHC)  
356 variability (Buckley et al. 2019; Foukal and Lozier 2018; Robson et al. 2012; Yeager and Robson 2017), as an  
357 important part of the Atlantic Meridional Overturning Circulation (AMOC; Klavans et al. 2019; Latif et al. 2006; Yan  
358 et al. 2018; Zhang 2017; Zhang and Zhang, 2015). The inclusion of mesoscale ocean features in HR results in greater  
359 decadal variability in the North Atlantic SST, but without a substantial increase in decadal predictability, except in the  
360 Gulf Stream and its extension. Slightly decreased decadal SST predictability in the North Atlantic Subpolar Gyre in  
361 HR is in better agreement with observational estimates. We speculate that this is largely due to substantial different  
362 mean states between HR and LR (Fig. 8). Except for the Eastern Tropical Pacific, regions with increased decadal SST  
363 predictability based on HR compared with LR are in good agreement with regions with longer persistence of decadal-  
364 scale SST variability (Fig. 7d), suggesting that the paradox is less likely to occur in the HR model, based on the  
365 Markov model framework (see Eq. 5 in Section 2c).

366



367

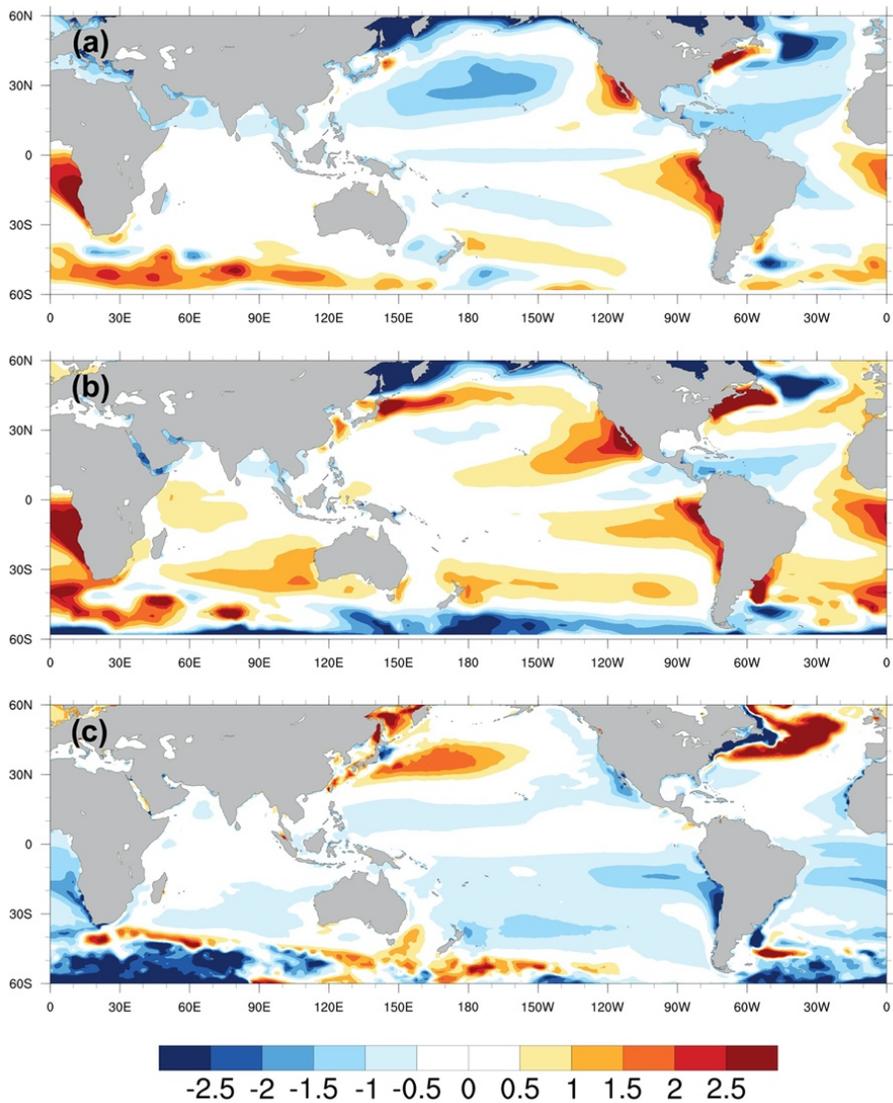
368

369

370

371

Fig. 7. Decadal SST predictability using CCSM4 models and the effect of ocean model resolutions. (a) Decadal SST predictability based on HR eddy-resolving CCSM4. Ratio of HR and LR CCSM4 in terms of (b) decadal SST predictability, (c) decadal SST variability, and (d) the persistence of decadal SST variability. The persistence of the system is estimated as the lag-1 year autocorrelation of 5-year low-pass filtered SST data.



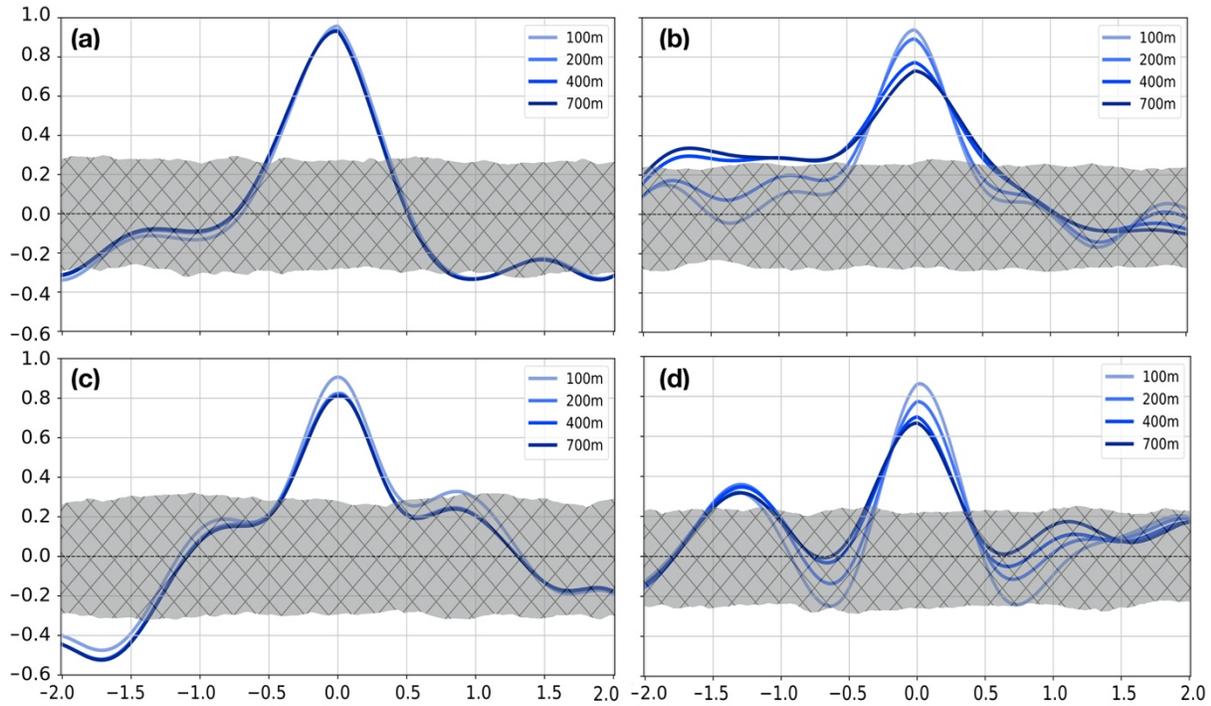
372

373 Figure 8. Annual mean SST bias in (a) 30 CMIP5 models and (b) CCSM4 based on HIST simulations. (c) Difference  
 374 of annual mean SST in HR and LR CCSM4 simulations. The unit is °C.

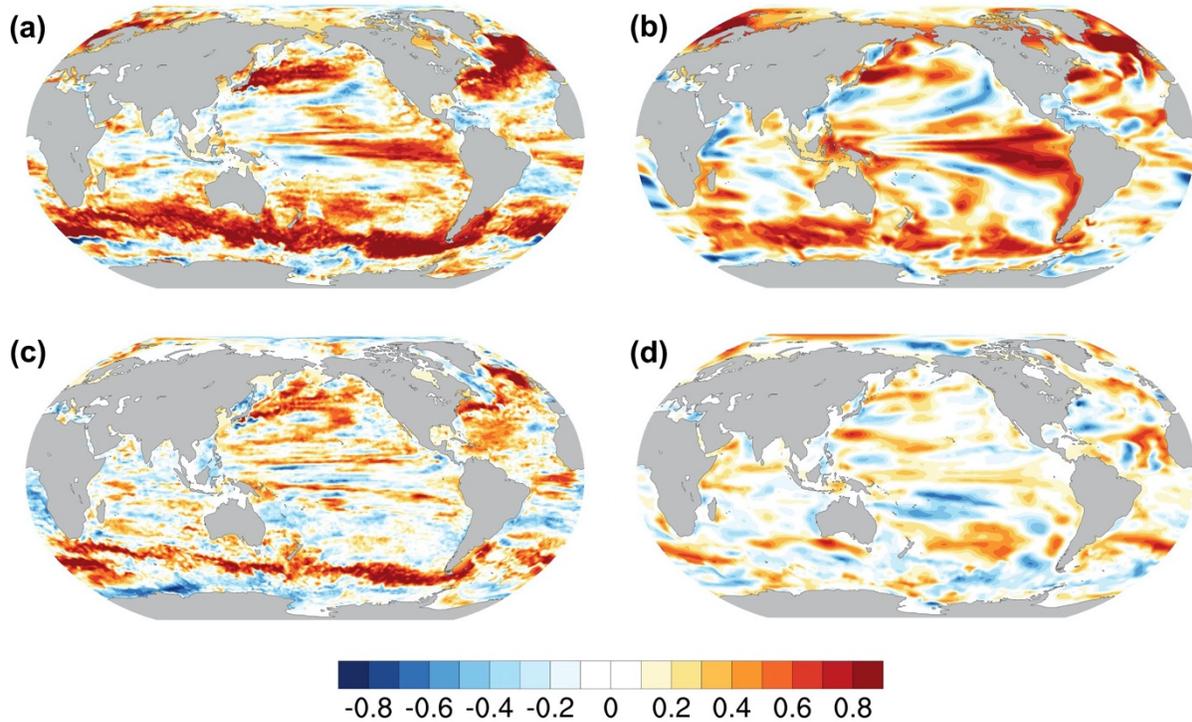
375

376 As noted early, we hypothesize on these decadal time-scales (5-10 years) the enhanced vertical connectivity  
 377 in the HR simulation compared to the LR simulation is a possible explanation for the longer limits of predictability  
 378 (Buckley et al. 2019). We demonstrate this point by taking a close look at the vertical structure over the Gulf Stream  
 379 and eastern Tropical Atlantic regions where large differences in decadal SST predictability are situated (Fig. 7b). For  
 380 shallow depth integrals of 100m and 200m, both LR and HR models show OHC anomalies highly correlated with the  
 381 SSTA, especially in the Gulf Stream (Figs. 9a and 9b). However, for deeper depth integrals, such as to 400m and

382 700m, the OHC anomalies get progressively less correlated with the SSTa in the LR model, especially in the eastern  
 383 Tropical Atlantic, indicating evident differences with HR, which maintain a consistent vertical structure.  
 384



385  
 386 Fig. 9. Lagged cross-correlation between SSTa and OHC anomalies for depths of integration to 100, 200, 400 and  
 387 700 m in the Gulf Stream (32°N-45°N; 80°W-45°W) region for (a) HR and (b) LR; (bottom) in Eastern Tropical  
 388 Atlantic (20°S-10°N; 15°W-15°E) for (c) HR and (d) LR. Negative (positive) years indicate the SSTa leading (lagging)  
 389 the OHC anomalies for lags between -2 and 2 years. All variables are 5-year low-pass filtered. The one-tailed (95%)  
 390 significance threshold for the cross-correlation is depicted by the hatched area and estimated using the non-parametric  
 391 random phase method (Ebisuzaki 1997).



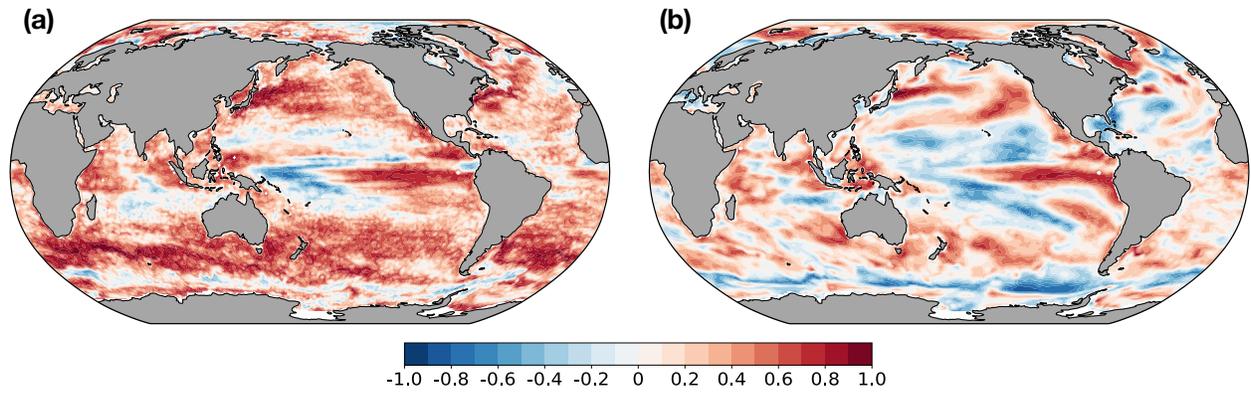
392

393 Fig. 10. Vertical correlation between SSTA and subsurface ocean temperature anomalies averaged over (a, b) the  
 394 upper 200-500 m and (c, d) the upper 700-1000 m. (a) and (c) are estimated based on HR, while (b) and (d) are  
 395 estimated based on LR. All the data are 5-year low-pass filtered.

396

397 To underscore the results in Fig. 9, i.e., the enhanced vertical connectivity associated with resolved ocean  
 398 meso-scale features and processes we show the vertical correlation between SSTA and subsurface ocean temperature  
 399 anomalies (Fig. 10) and the instantaneous correlation between the 400m OHC anomalies and surface heat flux  
 400 anomalies (Fig. 11). Fig. 10 shows the vertical correlation between SSTA and thermocline ocean temperature averaged  
 401 over 200-500 m (upper thermocline; Figs. 10a and 10b) and 700-1000 m (deeper thermocline; Figs. 10c and 10d).  
 402 We note that there is strong vertical connection between SST and upper thermocline in both HR (Fig. 10a) and LR  
 403 (Fig. 10b). The most significant difference between HR and LR occurs in the deeper thermocline. Much stronger  
 404 vertical connectivity is seen over eddy-rich regions in HR (Fig. 10c) compared with LR (Fig. 10d). Fig. 11 shows the  
 405 correlation between the surface net heat flux anomalies and OHC anomalies for HR (Fig. 11a) and for LR (Fig. 11b).  
 406 The results are generally consistent with Buckley et al. (2019). The correlations in frontal ocean zones and eddy rich  
 407 regions in HR (Fig. 11a) are stronger in HR than LR (Fig. 11b). Further, in LR there is widespread regions of negative

408 correlations where in the Gulf Stream and Kuroshio subtropical recirculation gyres, subtropical North and South  
409 Pacific and Southern Ocean (Fig. 11b). This reversal is difficult to detect or appears only in isolated regions of HR,  
410 notably in the Pacific warm pool and the sub-Antarctic zone of deep mixed layers.



411  
412 Fig. 11. Correlation between 5yr low-pass filtered anomalies of 0-400m ocean heat content and net heat flux for (a)  
413 HR and (b) LR. The sign convention is positive heat flux warms the atmosphere.

414  
415 **4 Summary**

416 To understand the underlying mechanisms for the signal-to-noise paradox, we focus on two main questions:  
417 (i) where and to what extent is the paradox leading to substantial underestimates of the limit of predictability? (ii) Is  
418 this underestimate of predictability and associated signal-to-noise paradox related with the representation of ocean  
419 mesoscale processes and features? To address the first question, we re-examine decadal predictability in CMIP5  
420 models from the perspective of the signal-to-noise paradox. We first compare decadal predictability of SST and SLP  
421 in observations and CMIP5 models, showing that decadal predictability estimates based on models are generally  
422 underestimated, particularly in the Tropical Atlantic, Tropical Indian ocean, and many eddy-rich regions. The  
423 distribution of the signal-to-noise paradox in the SST fields of CMIP5 models is presented following the Markov  
424 model framework in Zhang and Kirtman (2019b). The difference between observed and model-simulated decadal  
425 predictability is closely associated with the signal-to-noise paradox in that models are likely to underestimate decadal  
426 predictability in regions where it is likely to have the signal-to-noise paradox. We also examined this question in the  
427 context of so-called historical climate simulations as well as pre-industrial control runs. For example, the Tropical  
428 Indian and Tropical Atlantic oceans are two typical regions significantly influenced by external forcing, where we  
429 detect lower chance of existence for the signal-to-noise paradox in CMIP5 HIST simulations compared with PI

430 simulations. Considerable regions in the North Atlantic also are impacted by external forcing in terms of decadal SST  
431 predictability and the signal-to-noise paradox.

432 To address the second question, we perform CCSM4 HR model experiments with resolved mesoscale ocean  
433 in comparison with the LR model, and the results are also discussed through the lens of the signal-to-noise paradox.  
434 The design of model experiments is based on the hypothesis that the presence of ocean mesoscale processes and  
435 features and the associated vertical connectivity impact decadal variability, predictability, and the signal-to-noise  
436 paradox. This is for the first time that the signal-to-noise issue has been addressed with eddy-resolving GCMs. Our  
437 argument here is generally consistent with Strommen and Palmer (2019) and Zhang and Kirtman (2019b) in that we  
438 attribute the low signal-to-noise ratio to the lack of persistence, which can be seen in decadal predictability estimates  
439 in Fig. 7. We further argue that the lack of persistence in climate models stems from the lack of vertical connectivity  
440 in the subsurface ocean between ocean mixed layer and thermocline. The extent to which the differences in vertical  
441 connectivity between the HR and LR models has been discussed from the subsurface vertical structure in terms of the  
442 correlation between SSTA and OHC anomalies, the correlation between SSTA and subsurface ocean temperature  
443 anomalies, as well as the correlation between OHC anomalies and surface heat flux anomalies. We have demonstrated  
444 that in the HR model with resolved mesoscale ocean, there is consistent upper ocean vertical structure and strong  
445 vertical connection in the subsurface ocean that is weaker or even absent in the LR model. The most significant  
446 difference of vertical connectivity between HR and LR occurs in the deep ocean (i.e., deep thermocline). The  
447 differences of vertical connectivity in the HR and LR models can thus contribute to the differences of the persistence  
448 of decadal SST variability and the decadal SST predictability. We argue that the HR models with resolved mesoscale  
449 ocean may potentially (at least partially) eliminate the signal-to-noise issue and thus improve decadal-scale climate  
450 predictability.

451

452

453

454

455

456

457 **References**

- 458 Allan, R., & Ansell, T. (2006). A new globally complete monthly historical gridded mean sea level pressure dataset  
459 (HadSLP2): 1850–2004. *Journal of Climate*, 19(22), 5816–5842. <https://doi.org/10.1175/JCLI3937.1>
- 460 Boer, G. J. (2004). Long time-scale potential predictability in an ensemble of coupled climate models. *Climate*  
461 *dynamics*, 23(1), 29–44. <https://doi.org/10.1007/s00382-004-0419-8>
- 462 Bryan, F. O., Tomas, R., Dennis, J. M., Chelton, D. B., Loeb, N. G., & McClean, J. L. (2010). Frontal scale air–sea  
463 interaction in high-resolution coupled climate models. *Journal of Climate*, 23(23), 6277–6291.  
464 <https://doi.org/10.1175/2010JCLI3665.1>
- 465 Buckley, M. W., DelSole, T., Lozier, M. S., & Li, L. (2019). Predictability of North Atlantic Sea Surface Temperature  
466 and Upper-Ocean Heat Content. *Journal of Climate*, 32(10), 3005–3023. [https://doi.org/10.1175/JCLI-D-18-](https://doi.org/10.1175/JCLI-D-18-0509.1)  
467 [0509.1](https://doi.org/10.1175/JCLI-D-18-0509.1)
- 468 Clement, A., K. Bellomo, L. N. Murphy, M. A. Cane, T. Mauritsen, G. Rädel, and B. Stevens (2015), The Atlantic  
469 Multidecadal Oscillation without a role for ocean circulation, *Science*, 350(6258), 320–324,  
470 [doi:10.1126/science.aab3980](https://doi.org/10.1126/science.aab3980). <https://doi.org/10.1126/science.aab3980>
- 471 Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., ... & Brönnimann, S. (2011). The  
472 twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, 137(654), 1–28.  
473 <https://doi.org/10.1002/qj.776>
- 474 Ding, R., Li, J., Zheng, F., Feng, J., & Liu, D. (2016). Estimating the limit of decadal-scale climate predictability using  
475 observational data. *Climate dynamics*, 46(5-6), 1563–1580. <https://doi.org/10.1007/s00382-015-2662-6>
- 476 Eade, R., Smith, D., Scaife, A., Wallace, E., Dunstone, N., Hermanson, L., & Robinson, N. (2014). Do seasonal-to-  
477 decadal climate predictions underestimate the predictability of the real world?. *Geophysical Research*  
478 *Letters*, 41(15), 5620–5628. <https://doi.org/10.1002/2014GL061146>
- 479 Ebisuzaki, W., 1997: A method to estimate the statistical significance of a correlation when the data are serially  
480 correlated. *J. Climate*, 10(9), 2147–2153. [https://doi.org/10.1175/1520-0442\(1997\)010<2147:AMTETS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1997)010<2147:AMTETS>2.0.CO;2)
- 481 Foukal, N. P., & Lozier, M. S. (2018). Examining the origins of ocean heat content variability in the eastern North  
482 Atlantic subpolar gyre. *Geophysical Research Letters*, 45(20), 11–275. <https://doi.org/10.1029/2018GL079122>
- 483 Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., et al. (2011). The community  
484 climate system model version 4. *Journal of Climate*, 24(19), 4973–4991. <https://doi.org/10.1175/2011JCLI4083.1>

485 Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., ... & Kirtman, B. P. (2013). A verification  
486 framework for interannual-to-decadal predictions experiments. *Climate Dynamics*, 40(1-2), 245-272.  
487 <https://doi.org/10.1007/s00382-012-1481-2>

488 Gonzalez, P. L., & Goddard, L. (2016). Long-lead ENSO predictability from CMIP5 decadal hindcasts. *Climate*  
489 *Dynamics*, 46(9-10), 3127-3147. <https://doi.org/10.1007/s00382-015-2757-0>

490 Guemas, V., Corti, S., García-Serrano, J., Doblas-Reyes, F. J., Balmaseda, M., & Magnusson, L. (2013). The Indian  
491 Ocean: The region of highest skill worldwide in decadal climate prediction. *Journal of climate*, 26(3), 726-739.  
492 <https://doi.org/10.1175/JCLI-D-12-00049.1>

493 Gupta, A. S., Jourdain, N. C., Brown, J. N., & Monselesan, D. (2013). Climate drift in the CMIP5 models. *Journal of*  
494 *Climate*, 26(21), 8597-8615. <https://doi.org/10.1175/JCLI-D-12-00521.1>

495 Harlaß, J., Latif, M., & Park, W. (2018). Alleviating tropical Atlantic sector biases in the Kiel climate model by  
496 enhancing horizontal and vertical atmosphere model resolution: climatology and interannual variability. *Climate*  
497 *dynamics*, 50(7-8), 2605-2635. <https://doi.org/10.1007/s00382-017-3760-4>

498 He, J., Kirtman, B., Soden, B. J., Vecchi, G. A., Zhang, H., & Winton, M. (2018). Impact of ocean eddy resolution on  
499 the sensitivity of precipitation to CO<sub>2</sub> increase. *Geophysical Research Letters*, 45(14), 7194-7203.  
500 <https://doi.org/10.1029/2018GL078235>

501 Hirahara, S., Ishii, M., & Fukuda, Y. (2014). Centennial-scale sea surface temperature analysis and its  
502 uncertainty. *Journal of Climate*, 32(16). <https://doi.org/10.1175/JCLI-D-12-00837.1>

503 Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., ... & Zhang, H. M. (2017).  
504 Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and  
505 intercomparisons. *Journal of Climate*, 30(20), 8179-8205. <https://doi.org/10.1175/JCLI-D-16-0836.1>

506 Infanti, J. M., & Kirtman, B. P. (2019). A comparison of CCSM4 high-resolution and low-resolution predictions for  
507 south Florida and southeast United States drought. *Climate dynamics*, 52(11), 6877-6892.  
508 <https://doi.org/10.1007/s00382-018-4553-0>

509 IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth  
510 Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and  
511 L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

512 Keenlyside, N. S., Latif, M., Jungclauss, J., Kornbluh, L., & Roeckner, E. (2008). Advancing decadal-scale climate  
513 prediction in the North Atlantic sector. *Nature*, 453(7191), 84. <https://doi.org/10.1038/nature06921>

514 Kim, W. M., Yeager, S., Chang, P., & Danabasoglu, G. (2018). Low-frequency North Atlantic climate variability in  
515 the community earth system model large ensemble. *Journal of Climate*, 31(2), 787-813.  
516 <https://doi.org/10.1175/JCLI-D-17-0193.1>

517 Kirtman, B. P., & Schopf, P. S. (1998). Decadal variability in ENSO predictability and prediction. *Journal of*  
518 *Climate*, 11(11), 2804-2822. [https://doi.org/10.1175/1520-0442\(1998\)011<2804:DVIEPA>2.0.CO;2](https://doi.org/10.1175/1520-0442(1998)011<2804:DVIEPA>2.0.CO;2)

519 Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., ... & Stan, C. (2012). Impact of ocean model  
520 resolution on CCSM climate simulations. *Climate dynamics*, 39(6), 1303-1328. [https://doi.org/10.1007/s00382-](https://doi.org/10.1007/s00382-012-1500-3)  
521 [012-1500-3](https://doi.org/10.1007/s00382-012-1500-3)

522 Kirtman, B. P., Min, D., Infanti, J. M., Kinter III, J. L., Paolino, D. A., Zhang, Q., ... & Peng, P. (2014). The North  
523 American multimodel ensemble: phase-1 seasonal-to-interannual prediction; phase-2 toward developing  
524 intraseasonal prediction. *Bulletin of the American Meteorological Society*, 95(4), 585-601.  
525 <http://dx.doi.org/10.1175/BAMS-D-12-00050.1>

526 Kirtman, B. P., Pegion, K., & Kinter, S. M. (2005). Internal atmospheric dynamics and tropical Indo-Pacific climate  
527 variability. *Journal of the atmospheric sciences*, 62(7), 2220-2233. <https://doi.org/10.1175/JAS3449.1>

528 Kirtman, B. P., Perlin, N., & Siqueira, L. (2017). Ocean eddies and climate predictability. *Chaos: An Interdisciplinary*  
529 *Journal of Nonlinear Science*, 27(12), 126902. <https://doi.org/10.1063/1.4990034>

530 Kirtman, B., Power, S. B., Adedoyin, A. J., Boer, G. J., Bojariu, R., Camilloni, I., ... & Prather, M. (2013). Near-term  
531 climate change: projections and predictability. In Stocker, T. F. et al. (eds.) *Climate Change 2013: The Physical*  
532 *Science Basis. Contribution of Working Group I. to the Fifth Assessment Report of the Intergovernmental Panel*  
533 *on Climate Change* (Cambridge University Press, 2013).

534 Klavans, J. M., Clement, A. C., & Cane, M. A. (2019). Variable External Forcing Obscures the Weak Relationship  
535 between the NAO and North Atlantic Multidecadal SST Variability. *Journal of Climate*, 32(13), 3847-3864.  
536 <https://doi.org/10.1175/JCLI-D-18-0409.1>

537 Knight, J. R., Andrews, M. B., Smith, D. M., Arribas, A., Colman, A. W., Dunstone, N. J., ... & Scaife, A. A. (2014).  
538 Predictions of climate several years ahead using an improved decadal prediction system. *Journal of*  
539 *Climate*, 27(20), 7550-7567. <https://doi.org/10.1175/JCLI-D-14-00069.1>

540 Kravtsov, S. (2012). An empirical model of decadal ENSO variability. *Climate dynamics*, 39(9-10), 2377-2391.  
541 <https://doi.org/10.1007/s00382-012-1424-y>

542 Kravtsov, S. (2020). Dynamics and predictability of hemispheric-scale multidecadal climate variability in an  
543 observationally constrained mechanistic model. *Journal of Climate*, (2020). <https://doi.org/10.1175/JCLI-D-19->  
544 [0778.1](https://doi.org/10.1175/JCLI-D-19-0778.1)

545 Kushnir, Y., Scaife, A. A., Arritt, R., Balsamo, G., Boer, G., Doblas-Reyes, F., ... & Matei, D. (2019). Towards  
546 operational predictions of the near-term climate. *Nature Climate Change*, 9(2), 94-101.  
547 <https://doi.org/10.1038/s41558-018-0359-7>

548 Latif, M., Collins, M., Pohlmann, H., & Keenlyside, N. (2006). A review of predictability studies of Atlantic sector  
549 climate on decadal time scales. *Journal of Climate*, 19(23), 5971-5987. <https://doi.org/10.1175/JCLI3945.1>

550 Li, J., Sun, C., & Jin, F. F. (2013). NAO implicated as a predictor of Northern Hemisphere mean temperature  
551 multidecadal variability. *Geophysical research letters*, 40(20), 5497-5502. <https://doi.org/10.1002/2013GL057877>

552 Marzocchi, A., Hirschi, J. J. M., Holliday, N. P., Cunningham, S. A., Blaker, A. T., & Coward, A. C. (2015). The  
553 North Atlantic subpolar circulation in an eddy-resolving global ocean model. *Journal of Marine Systems*, 142,  
554 126-143. <https://doi.org/10.1016/j.jmarsys.2014.10.007>

555 Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., ... & Karspeck, A. (2014). Decadal  
556 climate prediction: an update from the trenches. *Bulletin of the American Meteorological Society*, 95(2), 243-267.  
557 <https://doi.org/10.1175/BAMS-D-12-00241.1>

558 Merryfield, W. J., Baehr, J., Batté, L., Becker, E. J., Butler, A. H., Coelho, C. A., ... & Ferranti, L. (2020). Current  
559 and emerging developments in subseasonal to decadal prediction. *Bulletin of the American Meteorological*  
560 *Society*, (2020). <https://doi.org/10.1175/BAMS-D-19-0037.1>

561 Minobe, S., Kuwano-Yoshida, A., Komori, N., Xie, S. P., & Small, R. J. (2008). Influence of the Gulf Stream on the  
562 troposphere. *Nature*, 452(7184), 206-209. <https://doi.org/10.1038/nature06690>

563 Murphy, L. N., Bellomo, K., Cane, M., & Clement, A. (2017). The role of historical forcings in simulating the  
564 observed Atlantic multidecadal oscillation. *Geophysical Research Letters*, 44(5), 2472-2480.  
565 <https://doi.org/10.1002/2016GL071337>

566 Newman, M. (2007). Interannual to decadal predictability of tropical and North Pacific sea surface  
567 temperatures. *Journal of climate*, 20(11), 2333-2356. <https://doi.org/10.1175/JCLI4165.1>

568 O'Reilly, C. H., Weisheimer, A., Woollings, T., Gray, L. J., & MacLeod, D. (2019). The importance of stratospheric  
569 initial conditions for winter North Atlantic Oscillation predictability and implications for the signal-to-noise  
570 paradox. *Quarterly Journal of the Royal Meteorological Society*, 145(718), 131-146.  
571 <https://doi.org/10.1002/qj.3413>

572 Patricola, C. M., Li, M., Xu, Z., Chang, P., Saravanan, R., & Hsieh, J. S. (2012). An investigation of tropical Atlantic  
573 bias in a high-resolution coupled regional climate model. *Climate Dynamics*, 39(9-10), 2443-2463.  
574 <https://doi.org/10.1007/s00382-012-1320-5>

575 Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., ... & Trémolet, Y. (2016). ERA-20C: An  
576 atmospheric reanalysis of the twentieth century. *Journal of Climate*, 29(11), 4083-4097.  
577 <https://doi.org/10.1175/JCLI-D-15-0556.1>

578 Rayner, N. A. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., ... & Kaplan, A. (2003).  
579 Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth  
580 century. *Journal of Geophysical Research: Atmospheres*, 108(D14). <https://doi.org/10.1029/2002JD002670>

581 Richter, I. (2015). *Climate model biases in the eastern tropical oceans: causes, impacts and ways forward*. Wiley  
582 *Interdisciplinary Reviews: Climate Change*, 6(3), 345-358. <https://doi.org/10.1002/wcc.338>

583 Robson, J. I., Sutton, R. T., & Smith, D. M. (2012). Initialized decadal predictions of the rapid warming of the North  
584 Atlantic Ocean in the mid 1990s. *Geophysical Research Letters*, 39(19). <https://doi.org/10.1002/2016GL070559>

585 Samanta, D., Karnauskas, K. B., Goodkin, N. F., Coats, S., Smerdon, J. E., & Zhang, L. (2018). Coupled model biases  
586 breed spurious low-frequency variability in the tropical Pacific Ocean. *Geophysical Research Letters*, 45(19), 10-  
587 609. <https://doi.org/10.1029/2012GL053370>

588 Scaife, A. A., & Smith, D. (2018). A signal-to-noise paradox in climate science. *npj Climate and Atmospheric*  
589 *Science*, 1(1), 28. <https://doi.org/10.1038/s41612-018-0038-4>

590 Scaife, A. A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R. T., Dunstone, N., ... & Hermanson, L. (2014).  
591 Skillful long-range prediction of European and North American winters. *Geophysical Research Letters*, 41(7),  
592 2514-2519. <https://doi.org/10.1002/2014GL059637>

593 Scaife, A. A., Camp, J., Comer, R., Davis, P., Dunstone, N., Gordon, M., ... & Roberts, M. (2019). Does increased  
594 atmospheric resolution improve seasonal climate predictions?. *Atmospheric Science Letters*.  
595 <https://doi.org/10.1002/asl.922>

596 Shaffrey, L. C., Hodson, D., Robson, J., Stevens, D. P., Hawkins, E., Polo, I., ... & Smith, D. (2017). Decadal  
597 predictions with the HiGEM high resolution global coupled climate model: description and basic  
598 evaluation. *Climate Dynamics*, 48(1-2), 297-311. <https://doi.org/10.1007/s00382-016-3075-x>

599 Siegert, S., Stephenson, D. B., Sansom, P. G., Scaife, A. A., Eade, R., & Arribas, A. (2016). A Bayesian framework  
600 for verification and recalibration of ensemble forecasts: How uncertain is NAO predictability?. *Journal of*  
601 *Climate*, 29(3), 995-1012. <https://doi.org/10.1175/JCLI-D-15-0196.1>

602 Siqueira, L., & Kirtman, B. P. (2016). Atlantic near-term climate variability and the role of a resolved Gulf  
603 Stream. *Geophysical Research Letters*, 43(8), 3964-3972. <https://doi.org/10.1002/2016GL068694>

604 Smith, D. M., Eade, R., Scaife, A. A., Caron, L. P., Danabasoglu, G., DelSole, T. M., ... & Kharin, V. (2019). Robust  
605 skill of decadal climate predictions. *npj Climate and Atmospheric Science*, 2(1), 13.  
606 <https://doi.org/10.1038/s41612-019-0071-y>

607 Strommen, K., & Palmer, T. N. (2019). Signal and noise in regime systems: a hypothesis on the predictability of the  
608 North Atlantic Oscillation. *Quarterly Journal of the Royal Meteorological Society*, 145(718), 147-163.  
609 <https://doi.org/10.1002/qj.3414>

610 Sun, C., Li, J., & Jin, F. F. (2015). A delayed oscillator model for the quasi-periodic multidecadal variability of the  
611 NAO. *Climate dynamics*, 45(7-8), 2083-2099. <https://doi.org/10.1007/s00382-014-2459-z>

612 Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of*  
613 *the American Meteorological Society*, 93(4), 485-498. <https://doi.org/10.1175/BAMS-D-11-00094.1>

614 Ting, M., Kushnir, Y., Seager, R., & Li, C. (2009). Forced and internal twentieth-century SST trends in the North  
615 Atlantic. *Journal of Climate*, 22(6), 1469-1481. <https://doi.org/10.1175/2008JCLI2561.1>

616 Wang, G., Dommenget, D., & Frauen, C. (2015). An evaluation of the CMIP3 and CMIP5 simulations in their skill  
617 of simulating the spatial structure of SST variability. *Climate dynamics*, 44(1-2), 95-114.  
618 <https://doi.org/10.1007/s00382-014-2154-0>

619 Wittenberg, A. T., Rosati, A., Delworth, T. L., Vecchi, G. A., & Zeng, F. (2014). ENSO modulation: Is it decadal  
620 predictable?. *Journal of Climate*, 27(7), 2667-2681. <https://doi.org/10.1175/JCLI-D-13-00577.1>

621 Xu, Z., Chang, P., Richter, I., & Tang, G. (2014). Diagnosing southeast tropical Atlantic SST and ocean circulation  
622 biases in the CMIP5 ensemble. *Climate dynamics*, 43(11), 3123-3145. <https://doi.org/10.1007/s00382-014-2247->

624 Yan, X., Zhang, R., & Knutson, T. R. (2018). Underestimated AMOC variability and implications for AMV and  
625 predictability in CMIP models. *Geophysical Research Letters*, 45(9), 4319-4328.  
626 <https://doi.org/10.1029/2018GL077378>

627 Yeager, S. G., & Robson, J. I. (2017). Recent progress in understanding and predicting Atlantic decadal climate  
628 variability. *Current Climate Change Reports*, 3(2), 112-127. <https://doi.org/10.1007/s40641-017-0064-z>

629 Zhang, J., & Zhang, R. (2015). On the evolution of Atlantic meridional overturning circulation fingerprint and  
630 implications for decadal predictability in the North Atlantic. *Geophysical Research Letters*, 42(13), 5419-5426.  
631 <https://doi.org/10.1002/2015GL064596>

632 Zhang, L., Delworth, T. L., & Jia, L. (2017). Diagnosis of decadal predictability of Southern Ocean sea surface  
633 temperature in the GFDL CM2. 1 model. *Journal of Climate*, 30(16), 6309-6328. [https://doi.org/10.1175/JCLI-D-](https://doi.org/10.1175/JCLI-D-16-0537.1)  
634 [16-0537.1](https://doi.org/10.1175/JCLI-D-16-0537.1)

635 Zhang, R. (2017). On the persistence and coherence of subpolar sea surface temperature and salinity anomalies  
636 associated with the Atlantic multidecadal variability. *Geophysical Research Letters*, 44(15), 7865-7875.  
637 <https://doi.org/10.1002/2017GL074342>

638 Zhang, W., & Kirtman, B. (2019a). Estimates of Decadal Climate Predictability from an Interactive Ensemble  
639 Model. *Geophysical Research Letters*, 46(6), 3387-3397. <https://doi.org/10.1029/2018GL081307>

640 Zhang, W., & Kirtman, B. (2019b). Understanding the Signal-to-noise Paradox with a Simple Markov  
641 Model. *Geophysical Research Letters*. <https://doi.org/10.1029/2019GL085159>

642

643

644